Reply to Referee #1

We appreciate the important comments made by the reviewer and we hope that our manuscript has improved.

> The authors compare surface fluxes and PBL heights from one MM5 simulation and one WRF Version 3.2.1 simulation with observations. Although nicely presented, the paper needs definitely major revisions (including additional simulations) before is can be published in GMD.
>
> The authors use for their comparison a version of WRF that is almost four years old - they could have used at least version 3.5. They also test only one PBL scheme for their inter-comparison and draw quite general conclusions from this single simulation for a very short episode. Besides of this, the used PBL scheme is considered as obsolete in WRF and labeled as ’to be removed in the future’.
>
> Therefore, I think it is absolutely necessary to redo the study with a more recent version of WRF and also test more PBL schemes in order to be of more general interest and thus justify publication in GMD. Due the very short episode that is studied, this should be easily possible.

Answer:

In the revised version we performed new simulations with WRF version 3.5.1. We now used YSU as the PBL scheme in WRF and added the following sentence in Chapter 2.3 "WRF model":

"In past air quality studies in Houston we used YSU and found promising results (Czader et al., 2013), and recent intercomparisons with other PBL schemes for the same area showed that YSU simulates vertical meteorological profiles as satisfactorily as the Asymmetric Convective Model version 2 (ACM2); the Mellor–Yamada–Janjic (MYJ) and Quasi-Normal Scale Elimination (QNSE), but may be the best to replicate vertical mixing of ozone precurors (Cuchiara et al., 2014)."

Reply to Referee #2

We appreciate the important comments made by the reviewer and we hope that our manuscript has improved.

> The authors compare simulation results of MM5 and WRF for a grid cell of Houston (Texas, USA). They focus on temperature and wind as well as surface energy fluxes and planetary boundary layer height.
> In general, the discussed topic to study and review model development is useful to the community. However, the present paper shows several fundamental issues which require new simulations. Thus, I recommend a rejection of the paper in the current form.
>
> Issues:
> 1) The model versions of both the MM5 and WRF are old. Since the development of MM5 has been discontinued, why has the newest version of MM5 from 2006 not been used? Also the used WRF version is from 2010 and several newer versions are published!
> Furthermore, the authors state that there is no cumulus parametrization in WRF. This is not true: http://www2.mmm.ucar.edu/wrf/users/wrfv3.5/phys_references.html#CU In particular, the Grell and Devenyi Scheme used in MM5 is available.

Answer:

We performed new simulations with WRF version 3.5.1..

We did not perform additional simulations with the newest version of MM5 from 2006 for two reasons:

(i) as mentioned in the text we used MM5 for a baseline comparison, and the MM5 version 3.6.1 was used in the Houston area until recently (Ngan et al., 2012).

Ngan, F., D. Byun, H. C. Kim, D. G. Lee, B. Rappenglück, and A. Pour-Biazar, 2012. Performance assessment of retrospective meteorological inputs for use in air quality modeling during TexAQS 2006. Atmospheric Environment, 54: 86-96.

(ii) as MM5 has been discontinued we do not see any additional benefit to use a MM5 version from 2006.

We apologize for the confusion with regard to the cumulus parameterization in WRF. The cumulus setting is that cumulus is set to 0 (none) for the 4-km domain but there are cumulus schemes for 12 and 36-km. We added corresponding information.

> 2) Only one grid cell is used for the evaluation. The effective resolution of RCMs is several grid cells (Grasso 2000). Please justify that only one grid cell is used. Furthermore, how do the surface parameters in the model and in reality fit together? Which area is seen by the sensors and how does this correlate to the model parameters.

Answer:

For WRF and MM5 modeling more than one grid cell was used. As mentioned in the text, the horizontal grid scales were the 36-km CONUS domain, 12-km eastern Texas domain, and the 4-km Houston-Galveston-Brazoria domain. However, for evaluation only, simulation results are extracted from the 4-km domain grid cell centered over the UH Coastal Center and is thus a point-to grid cell comparison for that specific cell. We do not claim that these validations are valid throughout the domain and for each grid cell as this would require a corresponding network of micrometeorological observations. However, we believe that a point-to grid validation on one 4-km domain grid cell may still be helpful in elucidating different behaviours and/or progresses in different models to simulate boundary layer properties. We included corresponding statement in the conclusions.

With regard to the footprint of the observations and the correlation to the model parameter we added following sentences:

Chapter 2.1 Location:

"Most of the measurements used in this study were taken from 10 m, 2 m, or at the surface. Using a parameterized Langrangian back trajectory footprint model according to Kljun et al. (2004) it is possible to estimate maximum impact distances and 90% impact boundaries of the footprint affecting the observations (available at http://footprint.kljun.net/). Typical values for the daytime surface friction velocity u* are about 0.3-0.5 m/s, for the standard deviation of the vertical velocity fluctuations about 0.7-0.9 m/s. The roughness length is estimated to be 0.05

m. For 10 m measurements this yielded maximum impact distances of 85-100 m and 90% impact boundaries of 230-285 m, which is well within the surrounding prairie grass area."

Chapter 2.3 WRF model:
"All simulation results are taken from the 4-km domain grid cell centered over the UH Coastal Center (Figure 1) and is thus a point-to-grid cell comparison. The 90% boundaries of the footprint of the observations fall within this grid cell."

3) SI or SI-derived units have to be used (e.g. acres)

Answer:
We added the corresponding information in km^2.

4) While the definition of bias should be known to a wider audience, a short explanation of the r^2 would be appropriate. Furthermore, the r^2 only describes how large unsystematic errors are. Systematic errors are not visible due to the linear model. Thus, I propose to evaluate the model in terms of RMSE as well as systematic and unsystematic RMSE (Willmott 1981).

Answer:
We added some short explanation for r^2. We added RMSE data and evaluation.

5) Is the output precision of the water vapour of MM5 really so low that other values of WRF and measurements have to be rounded?

Answer:
This has been remedied. Now WRF and MM5 have the same precision (i.e. two digits).

6) Even if outgoing longwave radiation is not available as output, it can be calculate either from incoming longwave radiation, surface emissivity and temperature or from the energy balance at the surface of all other analysed fluxes!

Answer:

While we basically agree with the reviewer, our intention was to test model performances with their available outputs.

7) What do you mean with page 2718 line 16–18? What are other sources of shortwave radiation? The following description about the "NaN" is technical and does not belong in such a scientific paper.

Answer:

We mean that the models treat outgoing radiation as a direct decrease caused by albedo. We shortened the sentence accordingly. There was no intention to refer to any other source of shortwave radiation. We apologize for any confusion.

We replaced "NaN" by "-".

In particular the first two issues have to be clarified before a more detailed analysis is useful.